DAMAGE DETECTION IN MECHANICAL STRUCTURES USING EXTREME VALUE STATISTICS

Keith Worden*, David Allen**, Hoon Sohn**, Charles R. Farrar**
*Dynamics Research Group, Department of Mechanical Engineering,
University of Sheffield, Sheffield, UK.
**Engineering Sciences and Applications Division, Weapons Response Group,
Los Alamos National Laboratory, New Mexico, USA.

ABSTRACT

The first and most important objective of any damage identification algorithms is to ascertain with confidence if damage is present or not. Many methods have been proposed for damage detection based on ideas of novelty detection founded in pattern recognition and multivariate statistics. The philosophy of novelty detection is simple. Features are first extracted from a baseline system to be monitored, and subsequent data are then compared to see if the new features are *outliers*, which significantly depart from the rest of population. In damage diagnosis problems, the assumption is that outliers are generated from a damaged condition of the monitored system. This damage classification necessitates the establishment of a decision boundary. Choosing this threshold value is often based on the assumption that the parent distribution of data is Gaussian in nature. While the problem of novelty detection focuses attention on the outlier or extreme values of the data *i.e.* those points in the tails of the distribution, the threshold selection using the normality assumption weighs the central population of data. Therefore, this normality assumption might impose potentially misleading behavior on damage classification, and is likely to lead the damage diagnosis astray. In this paper, *extreme value statistics* is integrated with the novelty detection to specifically model the tails of the distribution of interest. Finally, the proposed technique is demonstrated on simulated numerical data and time series data measured from an eight degree-of-freedom spring-mass system.

Keywords: Extreme value statistics, novelty detection, damage detection, time series analysis, vibration test

1. INTRODUCTION

This paper is concerned with *novelty detection* in an unsupervised learning mode, which is the first level of damage identification. When applied to structural health monitoring, unsupervised learning means that data from the damaged condition are not available to aid in the damage detection process. The objective of unsupervised novelty detection is to establish a model of the system or structure's normal condition and thereafter to signal significant departures from this condition. In many ways, the technology of novelty detection encompasses traditional condition monitoring. However, the new term is a convenient means of recognizing the significant inputs to the field from multivariate statistics and pattern recognition that have recently occurred.

The first objective of novelty detection is to establish a model of the normal system condition based on the damage-sensitive features extracted from measured system response data. This objective can be accomplished in several ways. The more direct methods seek to model the probability distribution of the normal condition using *a priori* training data. One of the simplest, the outlier approach (Worden et al., 2000a) assumes a Gaussian distribution for the damage sensitive features and parameterizes the model distribution using estimated means and covariance matrices. More sophisticated approaches use Gaussian mixture models (Roberts, 1998 and 2000) or kernel density estimates (Worden et al., 2000b). The main limitation of all of these methods is that they make unwarranted assumptions about the nature of the feature distribution tails. These assumptions are potentially hazardous, as the extreme events that reside in the tails of the normal condition are likely to be misinterpreted. More specifically, novelty detection constructs a model based entirely on *central statistics* (the mean vector and covariance matrix) and the analysis is largely insensitive to the structure of the tails. Another way of regarding this problem is as a question of setting an appropriate threshold for novelty. If the true distribution of the structural normal condition is heavy tailed, this threshold will be too liberal and there are likely to be many false positives, indicating damage when the structure is in reality undamaged.

The major problems with modeling the undamaged condition of a system are that the functional form of the distribution is unknown and there that are an infinite number of candidate distributions that may be appropriate for

the prediction applications. Currently, a choice among the infinite distributions is made by a knowledgeable operator and then estimate parameters based on training data. This process is largely subjective. Any choice of distribution and parameters will also constrain the behavior of the tails to that prescribed distribution. Also in some cases, the extreme values of an event may be the only data that are recorded due to sensor or storage limitations so modeling the data as a parent distribution could also bring about erroneous results.

In fact, there is a large body of statistical theory that is explicitly concerned with modeling the tails of distributions, and these statistical procedures can be applied to the problem of novelty detection. The relevant field is referred to as *extreme value statistics (EVS)*, a branch of *order statistics*. There are many excellent textbooks and monographs in this field. Some are considered classics (Gumbel, 1958; Galambos, 1978), and others are more recent (Embrechts et al., 1997; Kotz and Nadarajah, 2000; Reiss and Thomas, 2001). Castillo (1988) is notable in its concern with engineering problems in fields like meteorology, hydrology, ocean engineering, pollution studies, strength of materials, etc. Although extreme value statistics has been widely applied, there has been little application of these techniques to novelty detection. Roberts introduced the ideas of extreme value statistics into novelty detection in (Roberts, 1998 and 2000) and applied them in the biosignal processing context. Although these studies broke new ground, they could be said to have fallen prey to making unwarranted assumptions. It was assumed that the distribution of the base data could be modeled adequately by a Gaussian mixture model. This assumption in turn forced the Mahanobis distances from the Gaussian centers into a chi-square distribution, which then forced a Gumbel form distribution on the extreme values. Gumbel is one distribution type out of three candidate extreme values distributions. This paper illustrates the use of extreme value statistics in their own right and not as another way of looking at Gaussian distributions in an effort to avoid such assumptions.

2. METHODOLOGY

2.1. Extreme Value Statistics

The Gaussian distribution occupies its central place in statistics for a number of reasons; not least is the central limit theorem (Benjamin and Cornell, 1970). The central limit theorem states that if $\{X_1, X_2, \ldots, X_n\}$ is a set of random variables with arbitrary distributions, the sum variable $X_\Sigma = X_1 + X_2 + \ldots + X_n$ will have a Gaussian distribution as $n \to \infty$. Although this theory is arguably the most important limiting theorem in statistics, it is not the only one. If the problem at hand is concerned with the tails of distributions, there is another theorem that is more appropriate.

Suppose that one is given a vector of samples $\{X_1, X_2, \ldots, X_n\}$ from an arbitrary *parent distribution*. The most relevant statistic for studying the tails of the parent distribution is the maximum operator, $\max(\{X_1, X_2, \ldots, X_n\})$, which selects the point of maximum value from the sample vector. Note that this statistic is relevant for the right tail of a univariate distribution only. For the left tail, the minimum should be used. The pivotal theorem of extreme value statistics (Fisher, 1928) states that in the limit as the number of vector samples tends to infinity, the induced distribution on the maxima of the samples can only take one of three forms: Gumbel, Weibull, or Frechet. The rest of this section will be concerned with elaborating on this fact.

If the values of the sequence $X_1, X_2, \ldots, X_n$ are arranged in ascending order, the rth element of this sequence $X_{r:n}$ is called the rth *order statistic*. The basic question, which now arises is, what are the distributions of the order statistics, in particular, the minimum, $X_{1:n}$, and the maximum, $X_{n:n}$.

Following Castillo (1988), let $m_n(x)$ be the number of samples for which $X_j \leq x$. Each time one chooses a value X_j from the sample, one is conducting a Bernoulli experiment, an experiment that has one of two outcomes, with a probability $F(x)$, the Cumulative Distribution Function (CDF), that $X_j \leq x$, and the complementary probability, $(1 - F(x))$, that $X_j > x$. The CDF of $m_n(x)$ is therefore a binomial distribution with $F^k(x)$ denoting the probability of success,

$$F_{m_n(x)}(r) = \text{Prob}[m_n(x) \leq r] = \sum_{k=0}^{r} \binom{n}{k} F^k(x)[1 - F(x)]^{n-k} \tag{1}$$

Now, because the event $\{X_{r:n} \leq x\}$ is basically the same as the event $\{m_n(x) \geq r\}$ so that, $\text{Prob}[X_{r:n} \leq x] = \text{Prob}[m_n(x) \geq r] = 1 - \text{Prob}[m_n(x) < r]$, and it follows that $F_{X_{r:n}}(x) = 1 - F_{m_n(x)}(r-1)$, or

$$F_{X_{r:n}}(x) = \text{Prob}[X_{r:n} \leq x] = \sum_{k=r}^{n} \binom{n}{k} F^k(x)[1 - F(x)]^{n-k} \tag{2}$$

If one is concerned with the maximum of the sample, the relevant order statistic is $X_{n:n}$ and the relevant distribution is,

$$F_{X_{n:n}}(x) = F^n(x) \tag{3}$$

If one is concerned with the minimum of the sample, the relevant order statistic is $X_{1:n}$ and the appropriate distribution is,

$$F_{X_{1:n}}(x) = 1 - [1 - F(x)]^n \tag{4}$$

Concentrating now on the maximum, let $n \to \infty$, then the limit distribution for the maximum will satisfy

$$\lim_{n \to \infty} F^n(x) = \begin{cases} 1 & \text{If } F(x) = 1 \\ 0 & \text{If } F(x) < 1 \end{cases} \tag{5}$$

This distribution doesn't make sense because a CDF is developed on the assumption that it is continuous, but here the limit is discontinuous. The way around this discontinuity is to normalize the independent variable with a sequence of constants ($x \to a_n + b_n x$) in such a way that,

$$\lim_{n \to \infty} F^n(a_n + b_n x) = H(x) \tag{6}$$

where $H(x)$ is a non-degenerate limit function. In fact, it is required that $H(x)$ be continuous. The situation for minima is similar: a sequence of normalizations is required such that,

$$\lim_{n \to \infty} 1 - [1 - F(c_n + d_n x)]^n = L(x) \tag{7}$$

and again $L(x)$ is a non-degenerate continuous limit function.

The fundamental theorem of extreme value statistics is stated in Fisher and Tippett (1928).

Theorem I. (Feasible limit distributions for maxima):
The only three types of non-degenerate distributions $H(x)$ satisfying Equation (6) are

$$\text{FRECHET}: \quad H_{1,\beta}(x) = \begin{cases} \exp\left[-\left(\dfrac{\delta}{x-\lambda}\right)^{\beta}\right] & \text{if } x \geq \lambda \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

$$\text{WEIBULL}: \quad H_{2,\beta}(x) = \begin{cases} 1 & \text{if } x \geq \lambda \\ \exp\left[-\left(\dfrac{\lambda - x}{\delta}\right)^{\beta}\right] & \text{otherwise} \end{cases} \tag{9}$$

$$\text{GUMBEL}: \quad H_{3,0}(x) = \exp\left[-\exp\left(-\dfrac{x-\lambda}{\delta}\right)\right] \quad -\infty < x < \infty \text{ and } \delta > 0 \tag{10}$$

Or, in the appropriate form for minima,

Theorem II. (Feasible limit distributions for minima)
The only three types of non-degenerate distributions $L(x)$ satisfying Equation (7) are

$$\text{FRECHET:} \quad L_{1,\beta}(x) = \begin{cases} 1 - \exp\left[-\left(\dfrac{\delta}{\lambda - x} \right)^{\beta} \right] & \text{if } x \le \lambda \\ 1 & \text{otherwise} \end{cases} \tag{11}$$

$$\text{WEIBULL:} \quad L_{2,\beta}(x) = \begin{cases} 0 & x \le \lambda \\ 1 - \exp\left[-\left(\dfrac{x - \lambda}{\delta} \right)^{\beta} \right] & x > \lambda \end{cases} \tag{12}$$

$$\text{GUMBEL:} \quad L_{3,0}(x) = 1 - \exp\left[-\exp\left(\dfrac{x - \lambda}{\delta} \right) \right] \quad -\infty < x < \infty \text{ and } \delta > 0 \tag{13}$$

where λ, α, and β are the model parameters, which should be estimated from the data.

Now given samples of maximum or minimum data from a number of n-point populations, it is possible to select an appropriate limit distribution and fit a parametric model to the data. It is also possible to fit a model to portions of the parent distribution's tails, as the distribution of the tails is equivalent to the appropriate extreme value distribution. Once the parametric model is obtained, it can be used to compute an effective threshold for novelty based on the true statistics of the data as opposed to statistics based on a blanket assumption of a Gaussian distribution.

2.2. Parameter Estimation

Having established the appropriate limit distribution, the next stage in the analysis is to estimate the parameters of the chosen distribution. The actual parameter estimation technique employed in this study only fits parameters to one canonical model form – the Gumbel distribution for minima. Therefore, if the data are distributed as maxima, the transformations $x \to -x$ and $\lambda \to -\lambda$ carry each maximum CDF into the corresponding minimum CDF at least as far as optimization is concerned.

Suppose the data have the Weibull distribution for minima, then the transformation $Y = \ln(X - \lambda)$, carries the Weibull distribution X into the Gumbel distribution Y with the following relations between the parameters,

$$\lambda_G = \ln(\delta_W) \text{ and } \delta_G = \frac{1}{\beta_W} \tag{14}$$

where the subscripts G and W denote Gumbel and Weibull distributions, respectively. This transformation requires an *a priori* estimate of λ_W, but this transformation can be obtained by optimizing the linearity of the empirical CDF plot in Weibull coordinates.

If the data have the Frechet distribution for minima, the transformation $Y = -\log(\lambda - X)$ carries the Frechet distribution X into the Gumbel distribution Y, with the following relations between the parameters,

$$\lambda_G = -\ln(\delta_F) \text{ and } \delta_G = \frac{1}{\beta_F} \tag{15}$$

where the subscript F denotes a Frechet distribution. Again the prior estimation of λ_F is required and this estimate can be obtained by maximizing the linearity of the empirical CDF plot in Frechet coordinates.

After transforming either the Weibull and Frechet distribution to the Gumbel distribution, the parameter estimation problem is reduced to fitting the data to the limit distribution of the form in Equation (13). The optimization estimates the parameters λ and δ, which minimize some error criterion. Note that because all distribution types are now transformed to a Gumbel distribution, the subscript G for the Gumbel distribution is omitted hereafter. The most straightforward error criterion is the weighted least-squares method, which seeks to minimize the following objective function G,

$$G = \sum_{i=1}^{q} w_i [p_i - L_{3,0}(x_i; \lambda, \delta)]^2 \qquad (16)$$

where the training data are the points on the empirical CDF $\{(x_i, p_i), i = 1, \ldots, q\}$ and p_i's are an appropriate choice of plotting positions. w_i's are a set of weights. There are various possibilities that Castillo (1988) recommends.

3. ANALYSIS AND RESULTS

The following two examples illustrate the use of EVS for setting novelty detection thresholds. Three data sets from different distributions are first generated to illustrate the difference between assuming a normal distribution and specifically modeling the extreme data. These numeric examples are then followed by an analysis of data from an eight-degree-of-freedom (8DOF) spring-mass system. The 8 DOF system is analyzed in an undamaged state and three different damaged states.

3.1. Numeric data

Simulated random signals from three different distributions are used to demonstrate the usefulness of the EVS without any assumptions of the parent distribution. In each example, the 99% confidence interval is computed based on the three following distributions:

1. The assumed true parent distribution

2. A best-fit normal distribution where the sample mean and standard deviation are estimated from the random data generated from the assumed parent distribution.

3. A extreme value distribution, the parameters of which are estimated from either the top or bottom fraction of the simulated random data.

Hereafter, the confidence interval estimation methods based on the above three distributions are referred to as Method 1, Method 2, and Method 3, respectively.

Setting a confidence interval on the parent distribution using either method 1 or 2 is fairly trivial. The lower and upper limits of the confidence interval is constructed by choosing a type I error threshold, α. This threshold is related to the percentage of false positive errors in your base line data. For example, an α=.05 will correspond to a limit in which the lower 5% of the data is considered a false positive. Conversely a value of α=.95 will correspond to an upper limit in which the top 5% of the data is considered a false positive. By using α=0.005 and α=0.995 and the inverse CDF for either the known distribution or a best fit normal distribution, limits which encompass 99% of the data can be obtained. The remaining 1% of the data is therefore being considered false positive indications.

For method 3, the lower and upper limits are estimated using the estimated parameters and a user defined type I error bound. Knowing that the extreme values will be modeled by a Gumbel distribution as in Equations (10) and (13), the following is formulated to estimate the lower limit of the confidence interval,

$$\text{Lower limit: } x_m = \lambda + \delta \ln(-\ln(1 - \alpha)) \qquad (17)$$

where x_m is the threshold, λ and δ are obtained from the Gumbel parameter estimation and α is the type I error bound. The upper limit of the confidence interval is similarly formulated,

$$\text{Upper limit: } x_M = \lambda - \delta \ln(-\ln(\alpha)) \qquad (18)$$

Some care must be taken in selecting the α value for method 3 in order to obtain limits that are comparable when applied to the parent distribution. In these numeric examples, the lower or upper 10% of the data are selected from the parent distribution to be modeled as the extrema. An α value for method 3 that will result in a 1% false positive error in the parent distribution needs to be selected. When examining the parent distribution, 0.5% of the data will be an outlier, which translates to 5% of the extreme data being outliers. Therefore, the appropriate type I error bounds for method 3 would be α=0.05 and α=0.95. This will allow the limits from all of the methods to be compared.

The three distributions chosen for this study are normal, lognormal and gamma distributions. The normal distribution will provide a sanity check to make sure that the EVS and the inverse CDF provide similar thresholds. The lognormal and the gamma distributions are both skewed and will provide an opportunity to dramatically

illustrate the shortcomings of the confidence interval estimation based on a normal assumption of the data. In Castillo (1988) it is shown that both the minimum and the maximum for the normal and lognormal distributions can be modeled with a Gumbel distribution, thereby reducing the effort of finding the best-fit distribution in this example. The gamma distribution has a Gumbel distributed maximum and a Weibull minimum. Distributions were created and analyzed of varying sample size from n=1000 to n=1e6.

Looking at the normally distributed data in Figure 1, it is seen that the thresholds obtained from Methods 1 and 3 are comparable. In the case of normally distributed data, Method 2 is the same as Method 1. Both the lognormal and the gamma distributions in Figures 2 and 3 show a close match between methods one and three compared to the large error in method two. Because both the lognormal and the gamma distributions contain only positive data points, the lower limit based on normality completely misses all of the minimum values. In all of the test cases the Gumbel and the actual upper limit are comparable. Tables 1, 2 and 3 summarize the results of the parameter estimation and number of outliers for three sets of data from each of the three distributions. Only the first 1,000 data points are graphed for illustrative purposes in Figures 1, 2 and 3.
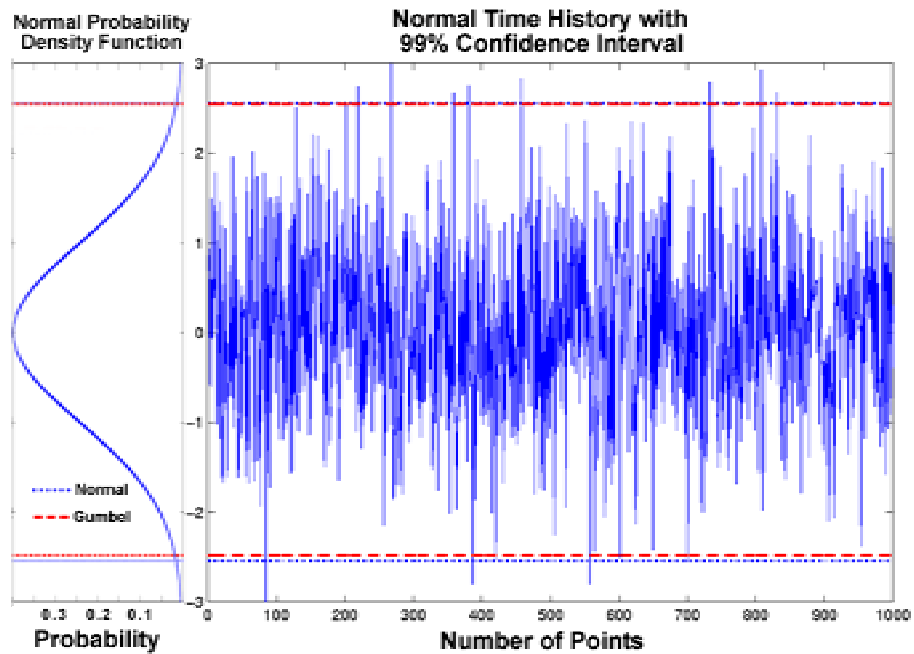

Figure 1 - The exact 99% confidence interval of a normal parent distribution compared with that from extreme values statistic. This figure shows the first 1000 data points from a 10,000 data point set.

The *least-squares return period relative error* (LSRPRE) method (Castillo, 1988) is used in fitting a Gumbel distribution to the maximum and minimum of the data. As can be seen in Table 1, even though method 3 returns thresholds that are slightly different from the known PDF, the number of outliers is more close to the expected 1%.

Table 1: Estimation of 99% confidence intervals for the 10,000 data points generated from a Gaussian parent distribution

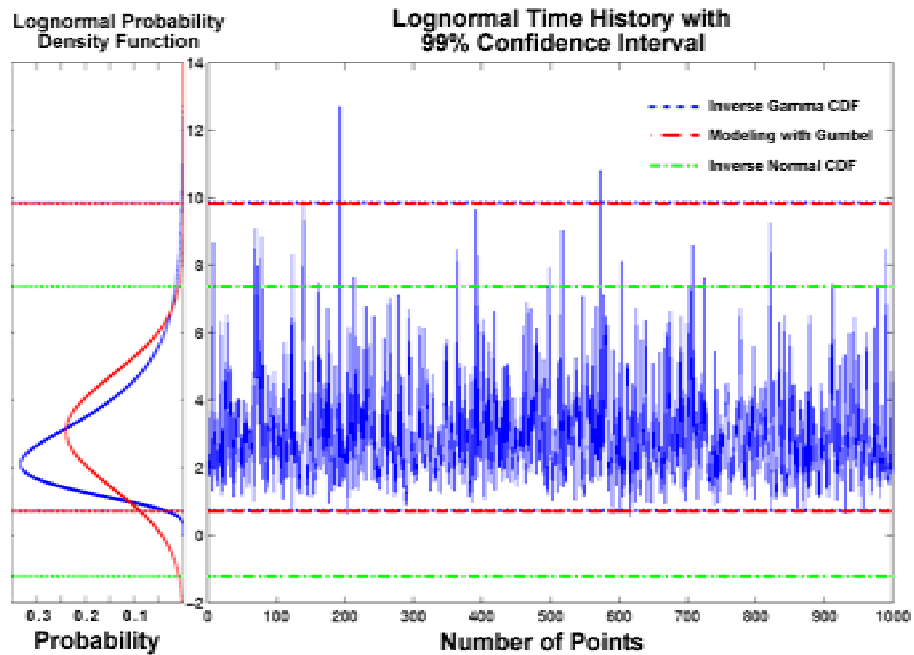| Estimation method	Upper confidence Limit	Lower confidence Limit	Number of outliers out of 10,000 samples
Method 1 (Expected)	2.548	-2.548	100
Method 2 (Normal)	2.551	-2.545	91
Method 3 (Gumbel)	2.549	-2.482	99

Figure 2 - The exact 99% confidence interval of a lognormal parent distribution compared with those computed from either extreme values statistic or the normality assumption.

Again the LSRPRE fitting method is employed for the maximum of the lognormal data. The minimum, however, is fitted using the *least-squares probability absolute error* method. In these numeric examples, several techniques of parameter estimation were employed from Castillo (1988) with the method which best fit the distribution being decided upon visually.

Table 2: Estimation of 99% confidence intervals for the 10,000 data points generated from a lognormal parent distribution

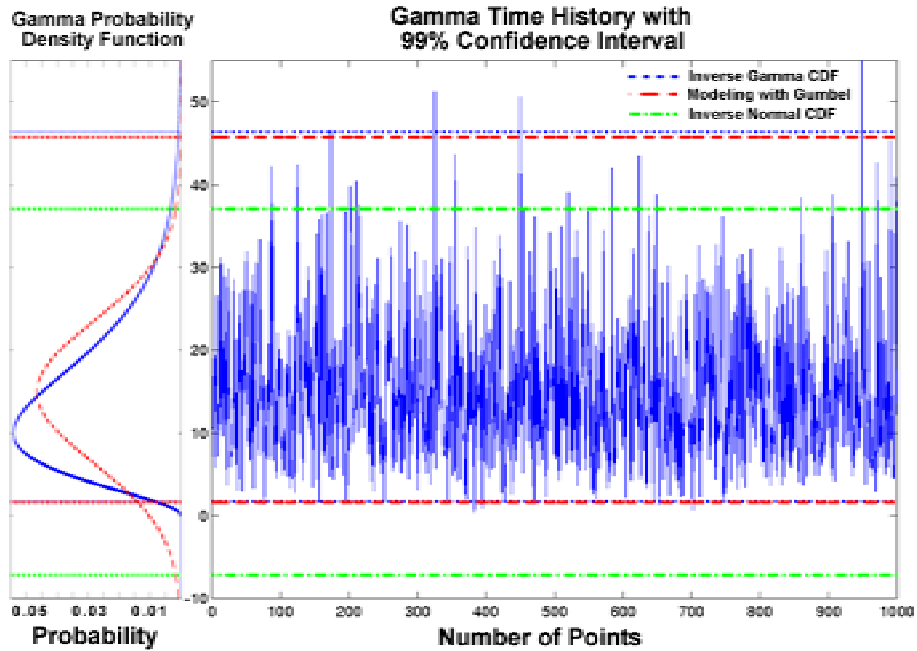| Estimation method	Upper confidence Limit	Lower confidence Limit	Number of outliers out of 10,000 samples
Method 1 (Expected)	9.854	0.750	100
Method 2 (Normal)	7.378	-1.206	230
Method 3 (Gumbel)	9.827	0.715	103

Figure 3 - The exact 99% confidence interval of a gamma parent distribution compared with those computed from either extreme values statistic or the normality assumption.

The maximum of the gamma parent distribution is fit using the LSRPRE method, while the minimum values are fit using *standard weighted least-squares* with a weighting factor of 1. The extreme value method again shows a distinct advantage over the normal assumption.

Table 3: Estimation of 99% confidence intervals for the 10,000 data points generated from a gamma parent distribution

Estimation method	Upper confidence Limit	Lower confidence Limit	Number of outliers out of 10,000 samples
Method 1 (Expected)	46.369	1.689	100
Method 2 (Normal)	37.016	-7.142	191
Method 3 (Gumbel)	45.693	1.600	96

A to the extreme values method is that a different method of parameter estimation is used to optimize the fit the minimum data in each of these examples. Once the extreme values are modeled well, there is, however, a noticeable advantage for novelty detection, even in the normally distributed example.

3.2. Eight degree-of-freedom spring-mass system

The effectiveness of the EVS is demonstrated using acceleration time series recorded from an 8 DOF spring mass system shown in Figure 4. The system is formed with eight translating masses connected by springs. Each mass is an aluminum disc of 25.4mm thick and 76.2mm in diameter with a center hole. The hole is lined with a Teflon bushing. There are small steel collars on each end of the discs (Figure 5). The masses all slide on a highly polished steel rod that supports the masses and constrains them to translate only along the rod. The masses are fastened together with coil springs epoxied to the collars that are, in turn, bolted to the masses.

The DOFs, springs and masses are numbered from the right end of the system, where the excitation is applied, to the left end as shown in Figure 4. The nominal value of mass 1 (m1) is 559.3 grams. Again, this mass is located at the right end where the shaker is attached. m1 is greater than the others because of the hardware needed to attach the shaker. All the other masses (m2 through m8) are 419.4 grams. The spring constant for all the springs is 56.7 kN/m for the initial condition. Damping in the system is caused primarily by Coulomb friction. Every effort is made to minimize the friction through careful alignment of the masses and springs. A common commercial lubricant is applied between the Teflon bushings and the support rod.

The undamaged configuration of the system is the state for which all springs are identical and have a linear spring constant. Nonlinear damage is defined as an occurrence of impact between two adjacent masses. Placing a bumper

between two adjacent masses so that the movement of one mass is limited relative to the other mass simulates damage. Figure 5 shows the hardware used to simulate nonlinear damage. When one end of a bumper, which is placed on one mass, hits the other mass, impact occurs. This impact simulates damage caused by the impact from the closing of a crack during vibration. Changing the amount of relative motion permitted before contact, and changing the hardness of the bumpers on the impactors can control the degree of damage. For all damage cases presented, the initial clearance is set to zero. Table 4 summarizes each of the four damage cases. In damage case 3, five of the twenty-five data sets were ignored because the excitation level was low enough that the bumpers did not contact the other mass, resulting in effectively undamaged cases.
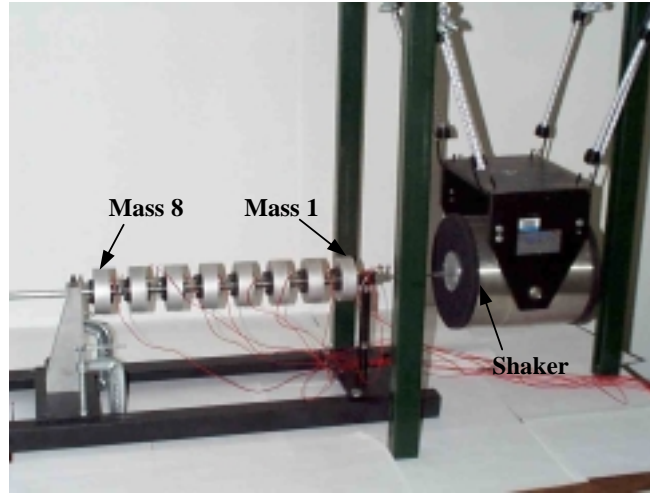

Figure 4 - An eight degrees-of-freedom system attached to a shaker with accelerometers mounted on each mass.
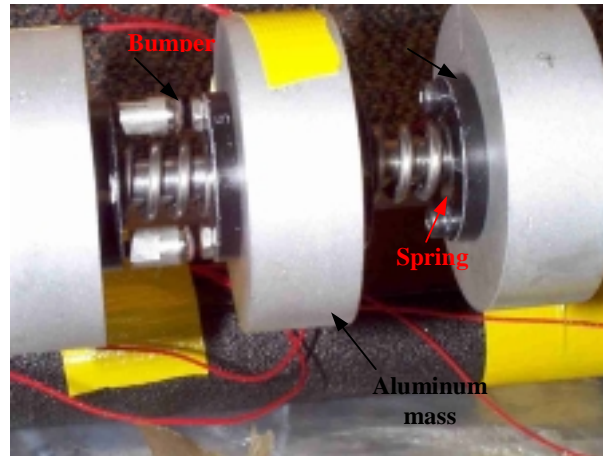

Figure 5 - A typical bumper used to simulate nonlinear damage.

Table 4 - List of time series employed in this study

Case	Description	Input level	Data # per input	Total data #
0	No bumper	3, 4, 5, 6, 7 Volts	15 sets	75 sets
1	Bumper between m1-m2	3, 4, 5, 6, 7 Volts	5 sets	25 sets
2	Bumper between m5-m6	3, 4, 5, 6, 7 Volts	5 sets	25 sets
3	Bumper between m7-m8	4, 5, 6, 7 Volts	5 sets	20 sets

In this example, a time series model called an AR-ARX model (Sohn et al, 2001) is first fit to an acceleration time history measured from the baseline condition of a system. If a time prediction model obtained from the baseline system is used to predict a new time signal measured under a damaged condition, the prediction errors will increase. Based on this premise, novelty detection is performed using the prediction errors as features. However, because the 8 DOF system is also subject to changing excitation levels, the varying input levels might result in unwanted false

outliers. To overcome this difficulty, an *auto-associative neural network* is employed for data normalization. Here, *data normalization* is a procedure of "normalizing" data sets such that signal changes caused by operational and environmental variations of the system can be separated from structural changes of interests, such as structural deterioration or degradation. Detailed discussion of data normalization using the auto-associative network can be found in Sohn et al., 2002.

After prediction errors are calculated, the undamaged features are analyzed to determine the threshold values for novelty detection. Typically in novelty detection one would construct a 'Novelty Index' (NI) by possibly subtracting the sample mean and dividing by sample standard deviation. A confidence interval is then set according to the baseline NI, and then subsequent sets of NI are tested against these threshold values much like the numeric examples. With the 8 DOF experiment a simple no NI is computed, rather a simple threshold crossing approach is used to determine the number of 'novel' prediction errors from each feature set. Statistical tools could then be applied to determine if the number of novel prediction errors related to a damaged state or not. This step is, however, omitted in this paper.

Because there are 4096 points in each case and a 99% confidence interval is being used, one would expect that for an undamaged case there would be 21 outliers on each side of the distribution, or 42 outliers in total. The outliers in the undamaged data were higher than was expected, but both the normal assumption and the extreme value method yielded similar results. Because several normality assessment techniques revealed that the prediction errors used as features were fairly close to normal, there was no surprise that the normality assumption worked as well as the extreme value statistics in this case. Table 5 summarizes the result of the 8 DOF experiment.

Table 5 - Summary of the 8 DOF system test results showing the predicted number of outliers contrasted with the normal assumption and the extreme value statistics. Highlighted cells represent the location of actual damage.

Number of outliers

Undamaged	Predicted	Normal	Gumbel		Damage 1	Predicted	Normal	Gumbel
m1	42	42	54		m1	42	47	60
m2	42	47	70		m2	42	381	369
m3	42	62	63		m3	42	123	118
m4	42	62	60		m4	42	93	90
m5	42	72	65		m5	42	89	78
m6	42	62	64		m6	42	63	88
m7	42	75	67		m7	42	73	68
m8	42	82	68		m8	42	93	73

Damage 2	Predicted	Normal	Gumbel		Damage 3	Predicted	Normal	Gumbel
m1	42	43	53		m1	42	41	54
m2	42	47	55		m2	42	50	58
m3	42	67	68		m3	42	78	77
m4	42	71	68		m4	42	71	64
m5	42	444	383		m5	42	103	88
m6	42	197	155		m6	42	**181**	**155**
m7	42	99	82		m7	42	586	526
m8	42	100	72		m8	42	331	290

** Highlighted cells show expected locations of increased outliers. Outliers shown in bold are potential false-positive indications of damage.

4. CONCLUSIONS

The results of this paper show that there are advantages of using extreme value statistics to analyze Structural Health Monitoring (SHM) data. Many of the SHM techniques implemented at Los Alamos National Laboratory (LANL) involve thresholds or decisions that are often based on a Gaussian distribution. The nature of novelty detection is to work with features in the extremities of a distribution that may not be accurately modeled by the Gaussian assumption. Despite limited scope, this paper shows positive results by reworking a simple threshold crossing technique to compute upper and lower limits on a confidence interval from extreme statistics instead of Gaussian statistics. These new limits based on extreme value statistics show a large improvement in the numeric examples. In the experimental example, the features were fairly close to normal, so both the Gaussian assumption and the extreme value statistics worked well. By incorporating extreme value statistics into damage identification and location algorithms, possibly limiting and erroneous assumptions in the routines can be removed, and distribution selection as well as parameter estimation is reduced.

ACKNOWLEDGEMENTS

REFERENCES

1. Benjamin J.R. & Cornell C.A., *Probability, Statistics and Decision for Civil Engineers*, McGraw-Hill, Inc. New York, NY, 1970.
2. Castillo E., *Extreme Value Theory in Engineering*, Academic Press Series in Statistical Modeling and Decision Science, San Diego, CA, 1988.
3. Embrechts P., Kluppelberg C. & Mikosch T., *Modeling Extremal Events*, Springer-Verlag, New York, NY, 1997.
4. Fisher R.A. & Tippett L.H.C., *Limiting forms of the frequency distributions of the largest or smallest members of a sample*, Proceedings of the Cambridge Philosophical Society 24 pp.180-190, 1928.
5. Galambos J., *The Asymptotic Theory of Extreme Order Statistics*, John Wiley and Sons, New York, NY, 1978.
6. Gumbel E.J., *Statistics of Extremes*, Columbia University Press, New York, 1958.
7. Kotz S. & Nadarajah S., *Extreme Value Distributions*. Theory and Applications. Imperial College Press, 2000.
8. Matlab, *The Language of Technical Computing*, The Math Works Inc., 1998.
9. Pickands III J., *Statistical inference using extreme order statistics*, Annals of Statistics 3 pp.119-131, 1975.
10. Reiss R.-D. & Thomas M., *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*, Birkhauser Verlag, 2001.
11. Roberts S., *Novelty detection using extreme value statistics*, IEE Proceedings in Vision, Image and Signal Processing 146 pp.124-129., 1998.
12. Roberts S., *Extreme value statistics for novelty detection in biomedical signal processing*, IEE Proceedings in Science, Technology and Measurement 147 pp.363-367, 2000
13. Sohn H. & Farrar C.R., "Damage diagnosis using time series analysis of vibration signals", *Smart Materials and Structures*, 10, pp. 446-451, 2001.
14. Sohn, H., Worden, K, and Farrar, C.R., "Consideration of Environmental and Operation Variations for Damage Diagnosis", *SPIE's 9th Annual International Symposium on Smart Structures and Materials,* San Diego, CA, USA, March 17-21, 2002.
15. Worden K., Manson G. & Fieller N.J., "Damage detection using outlier analysis", *Journal of Sound and Vibration*, 229 pp.647-667., 2000.
16. Worden K., Pierce S.G., Manson G., Philip W.R., Staszewski W.J. & Culshaw B., Detection of defects in composite plates using Lamb waves and novelty detection, International Journal of System Science 31 pp.1397-1409, 2000.